

## Comments on the ECOFRAM Terrestrial Draft Report

By Douglas J. Urban  
Senior Scientist  
Environmental Fate & Effects Division  
Office of Pesticide Programs  
U. S. Environmental Protection Agency

Date: June 11, 1999

### Approach to Comments

My comments will be divided into three parts: I) General comments, overview in nature, II) Comments on the tools, methods and processes proposed for probabilistic risk assessment in each section of the draft report; and, III) Response to questions identified in the "Charge to the Workshop Panel Members". Paragraphs under I) and II) will be numbered sequentially for easy reference.

### General Comments

1. The report is an excellent first step in developing a process, tools and methods for predicting the magnitude and probabilities of adverse effects to non-target terrestrial species resulting from the introduction of pesticides into the environment within the context of the FIFRA regulatory perspective and following the outline provided by the Framework for Ecological Risk Assessment (U.S. EPA, 1992) <sup>1</sup>

2. The ultimate goal of this initiative was to develop and validate risk assessment processes, tools and methods that address increasing levels of biological organization for individuals to populations, communities, and ecosystems, and account for direct and indirect effects. However, due to resource limitations, i.e., available data and time, the workgroup wisely (in my opinion) chose to focus on what was "doable". In Section 7.6.2 of the draft report, titled "Evaluation of How the Workgroup Fulfilled the Charge", the Workgroup provided an assessment on how well they fulfilled the charge. Their assessment was favorable, and generally I agree. The enormity and complexity of the ultimate goal of the initiative was daunting, and the report takes the Office of Pesticide Programs to the brink of probabilistic risk assessments for pesticides.

3. The members chose to limit the scope of their efforts by focusing on birds, on oral exposure, and on direct acute effects. Only limited attention was given to chronic effects. Impacts on other terrestrial taxa (such as small mammals and amphibians), other routes of exposure (such as dermal and inhalation), and other types of effects (such as indirect and sub-lethal) were recognized as being important but beyond the scope of this effort. The workgroup expected that the concepts and approach developed for the limited scope would be applicable to other taxa, other routes of exposure, and other types of effects. They believed that the recommendations could

---

<sup>1</sup> Adapted from the "Charge to the Terrestrial and Aquatic Workgroups, May 6, 1997.

serve as a model for future improvements.

#### Comments on the Methods & Tools and Process in Each Section

4. The report was divided into nine major sections. Section 8 listed references and Section 9 included 17 appendices. I will provide some comments on most sections but I will focus on the Effects Assessment, Section 4.0. My colleague, Dr. Edward Odenkirchen will focus his comments on the Exposure Assessment, Section 3.0. When I make a critical comment, I will always make a recommendation for change.

#### *Methods & Tools - Introduction*

5. Page 1-15. Since the choice of distribution shape can have a sizable effect on the risk analysis, especially in the tails, the workgroup recommended that risk assessors become familiar with the various methods for estimating distribution shapes and their limitations. I agree. There is a great need for in-house training for EFED scientists especially in statistical techniques for data analysis, interpretation, and presentation. As noted by the workgroup members, risk assessors need to become familiar with methods for estimating distribution shapes and their limitations.

6. Page 1-16. Implementation of probabilistic approaches to risk assessment will require many changes in the risk assessment process. Top among these, according to the workgroup, is the need for additional supporting data to reduce uncertainty in predicting effects. On page 1-18, the workgroup states that “large portion of the discussion in this report address the limitations in the available data and suggest ways to estimate or collect additional data to reduce the associated uncertainty.” I believe that the purpose and goal of probabilistic risk assessment should be to do just that, reduce uncertainties inherent in our risk assessment. This will take data beyond what we currently require for the registration and reregistration of pesticides. While the workgroup recommends methods and tools to implement probabilistic risk assessments using existing data, I believe that these simply will reveal the great uncertainty that exists with such estimates due to current data limitations. Current low-certainty results will do little to improve current decision making for pesticides over current estimates using deterministic quotients. When refinements are needed to reduce the uncertainty in predicted effects, they will require additional data. The greatest challenge will be to convince both regulatory divisions (risk managers) and the registrants to request and to pay for these additional data.

7. Page 1-18. The workgroup states that “conservatism is a value judgement deliberately introduced to account for uncertainty” and “risk managers need to understand the potential for distortions of the assessment due to cascading of biases from conservative assumptions.” As I will discuss later, the ‘a priori’ judgement that current deterministic quotients are conservative and replete with cascading biases

needs to be investigated. There are both effects and exposure data to suggest that they are not as conservative as we had thought.

The use of 5<sup>th</sup> percentile values for effects and dose in Level 1 screening assessments, as proposed by the workgroup, begins to account for some inherent variation in the current data. However, it also introduces a level of conservatism for less toxic pesticides that could result in decisions to progress to the higher 'probabilistic' Levels of Refinement. As a result, Level I could become largely dysfunctional.

I recommend that early in the implementation, efforts be made to establish exactly how conservative the screening level effects and exposure assessments really are. If conservatism is shown to be less than we assumed, then some different approaches to screening level assessments need to be found.

8. Page 1-19. The workgroup states that "not every assessment requires or warrants a quantitative estimate of the magnitude and probability of effects. Some may warrant it but it is impossible due to limitations in data and/or the understanding of the system requires assumptions which introduce such large uncertainties in the predicted effects that the assessments would not be scientifically defensible." Thus, the workgroup identified the need for screening tools "when data limitations imposed restrictions on full probabilistic techniques." However, the workgroup limited the utility of the screening tools to identifying minimal risk. They state that "if the inputs to the screening calculations have been established based on conservative assumptions, the certainty of the estimate of minimal risk should be, while maybe not quantified, relatively high. In cases where the potential for adverse effects is high along with a high level of certainty, further assessment may need to be considered."

9. Page 1-19. The above statement may be understood to conflict with page 6-9, Which Parts of the Assessment to Refine? In this section, the comparison of the risk prediction with the acceptability threshold provides a key to deciding how far to refine the assessment. Thus, further refinement may not be necessary if the initial prediction is either far enough above or below the threshold of acceptability. The example provided in Appendix C-10 is interesting. Figure 2. on page C10-6 of the appendix, shows that the preliminary deterministic point estimate is very far above the threshold of acceptability. Thus, it could be argued that, following the criteria on page 6-9, it would be more cost effective to seek mitigation measures than to invest in refining the assessment. However, the assessment continues. I question how far is 'far enough', and whether the third criteria, lines 23-24 on page 6-9 is really functional, especially considering the minimal risk limitation statements in section 1-19?

10. Page 1-22. The premise is that lower uncertainty requires additional information or data. Thus, probabilistic assessments based on current data will only be able to improve the quantification of the probability and magnitude of the risk; it will not be able to reduce the uncertainty inherent in the risk. The greatest need for

implementing probabilistic risk assessments, then, is **more data**. This should be stated more explicitly in each section of the report.

#### *Methods & Tools - Problem Formulation*

11. Page 2-4. While the risk managers acknowledged the limitations in the assessments, they indicated that the ecological risk assessments “should provide the most complete picture of ecological risks that are scientifically defensible.” Previously, the workgroup stated that large uncertainties in predicted effects would not be scientifically defensible. This continues to argue for additional data to reduce uncertainty.

12. Page 2-17. I agree that the three standard time scales, short, medium, and long, should be considered for both effects and exposure assessments, as well as for each level of refinement. Further, for screening Level 1 assessments, it seems logical that these time scales should generally follow the exposure periods used in the standard effect tests (currently <1 day, 5 days, 20 weeks for birds). One of the looming problems here is to match-up these exposure periods with reliable exposure data.

#### *Methods & Tools - Exposure Assessment*

13. Page 3-1. I agree that dose (expressed as mg/kg/day) better directly addresses the amount of chemical ingested by the individual that produces a response than residues on food items (expressed as ppm), and that estimating external dose is generally pragmatic. However, I would hope that we could find some estimates for birds, perhaps even chickens, of the differences between external and internal doses in order to provide some perspective and even some estimates of uncertainty for this assumption.

14. Page 3-10. Some suggestions where data for the various parameters and indices may be found would be helpful. The utility of these equations is significantly reduced when these are no sources of such information. The section becomes wonderful theory but little help in implementation.

15. Page 3-14 and C10-19. The biases that may affect the use of radio-tracking birds to estimate PT are identified on lines 14-17 page 3-16. However, no attempt is made to account for the uncertainty due to these biases in section 3.3.3 or under Model 6 in appendix C10. It seems quite important to do so at least in some situations since PT was determined to be quite influential in reducing risk in the example on page C10-17 and Figure 12.

16. Page 3-30. I agree with the recommendation to compile existing data on focal species into a single database “to facilitate future use of standard distributions by species and other significant sources of variation.” However, I believe that the

database needs to be thoroughly analyzed for certain trends and differences due to crop types, spatial and temporal differences. In this way, the risk assessor can select portions of the database that better fit specific use and exposure conditions.

17. Page 3-58. I agree that the ingestion, dermal and inhalation doses cannot be combined to arrive at an overall total dose. However, dermal and inhalation are major exposure pathways as indicated in Figure 3.1-1, and can contribute to the overall dose. Although the thickness of the arrows in the figure indicates the relative importance of the pathways, no further explanation or rationale is provided. I recommend that an expert panel or workshop be convened to provide some estimate of the importance of these pathways. If insufficient empirical data exists for the development of a distribution for use in probabilistic assessments, perhaps such a group could provide expert judgement that could be used to develop one pending the gathering of such data. [I make this recommendation numerous times in my comments to address situations where few data exist or where knowledge is lacking about factors that affect ecological risk assessment. The workgroup identified certain methods that would be helpful in these situations: subjective or Bayesian statistical methods, employing maximum entropy criteria to select distributions from a priori constraints, focusing on extreme value distributions when the tails are of interest, gathering empirically fitted distributions, and using default distributions such as the triangular or exponential (pages 1-12 to 1-15)].

If exposure distributions for dermal and inhalation could be developed and considering that they could not be combined with ingestion to arrive at a total dose, separate exposure assessments would be conducted for each. Decisions concerning the appropriate proportional contribution of each to the overall dose would be made during the risk assessment phase.

18. Pages 3-83 to -96. The workgroup identified a number of databases and models that could be used to generate estimates of pesticide residues for terrestrial exposure assessments. They all have potential biases that should be considered prior to using them as a basis for developing probabilistic distributions of potential residue levels on vegetation, insects and soil invertebrates. In general, I agree with the recommendations on pages 3-90 to -96. Implementing most of these will require substantial time and effort. However, I believe that the most straightforward way to address the lack of appropriate residue data in the interim for those pesticides where further refinement is necessary, i.e., the screening assessment indicates that there is a potential for adverse effects and the certainty of such effects occurring is high, is to require site specific residue data through FIFRA using 40 CFR part 158.75.

19. Page 3-111. The use of 5 or 10% exceedance values for use as high-end estimates of residue levels in invertebrates is appropriate; however, a significantly expanded insect residue database is needed prior to moving away from the theoretical surface area assumptions based on the nomograms derived from Kenaga and

Fletcher.

### *Methods & Tools - Effects Assessment*

20. Section 4.0 Some discussion is needed concerning how to deal with greater than values for LD50 tests. We have many in our data base. How should we (or should we?) use them in risk assessment? Currently, if the greater than value is >2000 mg/kg, we usually conclude no potential risk and do not attempt to use the result to calculate a quotient.

The in-house ecotoxicity data (OPP's EcoTox database) lacks the dose-response information. Thus, a near term recommendation for implementation of probabilistic tools for effects would be the re-capture of the specific dose-response information for existing studies in order to develop effects distributions based on empirical data sets.

21. Page 4-3. I agree that one of the greatest unknowns in the effects assessment is the relationship between laboratory results and effects in the field. The workgroup also identified other important unknowns that also contribute to uncertainty in effects assessments: the difference in inherent sensitivity between laboratory and field populations, the representativeness of the exposure scenario simulated in the laboratory, and the variable influence of stress of captivity on toxic responses among species. While none of these appear to be directly quantifiable, to simply ignore them will lessen the confidence in the results of the assessments and leave us with the same problems found in our current use of the quotient method. Thus, I recommend that an expert panel or workshop be convened to provide some estimate of the importance of these unknowns, how much they could contribute to the overall effects and risk assessment. Since they are unlikely to be quantified, perhaps such a group could provide expert judgement that could be used to develop a distribution for use in probabilistic assessments.

22. Page 4-6, 4-22 to -27. The workgroup identified indirect effects and sublethal effects as lacking adequate study, models and test data necessary to develop probabilistic assessment methods (page 4-22). It stated that additional work is required before these effects can be realistically addressed and noted that their proposed models and approaches do not and cannot address these effects. The workgroup did suggest that certain sublethal effects such as increased susceptibility to predation, could be included in any risk assessment by "incorporating a term for the anticipated loss of some individuals through decreased fitness." (page 4-24)

Based on the examples discussed in this section, indirect and sublethal effects can play a significant role in the effects assessment. Again, to simply ignore them will lessen the confidence in the results of the assessments and leave us with the same problems found in our current use of the quotient method. Thus, I recommend that an

expert panel or workshop be convened to provide some estimate of the importance of these effects, how much they could contribute to the overall effects and risk assessment. Pending collection of appropriate empirical data, perhaps such a group could provide expert judgement that could be used to develop distributions for use in probabilistic assessments.

If such distributions for indirect and sublethal effects could be developed it may not be appropriate to incorporate them into the effects assessment as a term for the anticipated loss of some individuals. Rather, as the workgroup indicated, each may need to be included in the risk assessment separately. Consideration of various sources of risk could be performed via some weight of evidence method or decision analysis.

23. Page 4-9. The workgroup noted that the dermal route of exposure can under certain circumstances be the dominate route. Since current testing is conducted through oral dosing and the standard risk assessment practices historically have not taken into account the other potential routes of exposure such as dermal, inhalation, the workgroup focused on the oral route for the establishing distributions of effects for probabilistic assessments.

I believe that effects assessments for dermal and inhalation exposure should be estimated. The importance of ocular exposure should also be determined. The workgroup argues that instead of dealing with problems with quantifying the exposure side of these exposures, more may be gained from modeling where dose is calculated as a body burden. In this case, the route of exposure is less important than the determinations of accumulated dose. However, determining the uncertainty inherent in mechanistic models and how well they predict real field situations is likely to present equally daunting problems. Again, I recommend that an expert panel or workshop be convened to provide some estimate of the importance of these effects, how much they could contribute to the overall effects and risk assessment. Pending the development of appropriate guidelines and data requirements including when required criteria, perhaps such a group could provide expert judgement that could be used to develop distributions for use in probabilistic assessments.

If effects distributions for dermal and inhalation could be developed and considering that they could not be combined with ingestion to arrive at a total dose, separate effects assessments would be conducted for each. Decisions concerning the appropriate proportional contribution of each route of exposure to the overall risk would be made during the risk assessment phase.

24. Page 4-17. Since the slope of the dose-response relationship is key to reducing the uncertainty in effects assessments, I would recommend that the ALD 'up-down' tests be modified to determine the full dose-response and the slope. I agree with the proposed changes in the avian acute oral LD50 test to reduce uncertainty at the

threshold dose, i.e., estimates of the LD5 or LD10.

25. Page 4-18 to -19. I agree that the avian LC50 test cannot be used as a quantitative descriptor of toxicity. I also agree with the proposed changes in study design.

26. Page 4-20 to -22. As the workgroup states, the avian reproduction test, is a rough screening tools for long-term effects. It is not designed to determine dose-response, and thus the probability of a specific magnitude of effect cannot be calculated. I agree that specific dose-response reproduction tests could be designed for pesticides where the potential for adverse effects is high along with a high level of certainty of effects and thus refinement in the risk assessment is needed.

I disagree, however, in the workgroup's conclusion that "uncertainties inherent in extrapolating from a laboratory reproduction study to reproductive effects of free-ranging birds with vastly different life histories are too great at this point to justify a major redesign of the current avian laboratory study to generate a dose-response relationship." (page 4-21) First of all, to simply say it's impossible to address the uncertainties in extrapolation will lessen the confidence in the results of the assessments and leave us with the same problems attendant to our current use of the quotient method. While there would be a problem extrapolating the results of testing precocial bobwhite quail and mallard ducks to altricial birds such as songbirds, extrapolating the results to other precocial bird species, especially other upland game birds and waterfowl would not be as difficult. Some expert guidance would be needed. Thus, I recommend that an expert panel or workshop be convened to provide some estimate of the importance of these extrapolations, how much they could contribute to the overall effects and risk assessment. Since they are unlikely to be quantified, perhaps such a group could provide expert judgement that could be used to develop a distribution for use in probabilistic assessments. Second, I recommend that a standardized reproduction protocol be developed for altricial birds. Third, I believe that a different approach to avian testing and test design may overcome design difficulties. Part of the problem encountered in designing avian reproduction tests to determine a dose-response is not knowing 'a priori' which of the many measured parameters will result in a response at what levels of exposure. This could be addressed by modifying the Level 2 for long term effects. Thus, if refinement is necessary after Level 1, an additional avian reproduction test would be required. A dose-response test would be designed based on the most sensitive response observed in the Level 1 test results. Alternately, the Level 1 avian reproduction test could be redesigned as an expanded range finding test. The test levels would be set based on maximum estimated environmental concentrations. While there would be additional costs with this approach, there will always be additional costs when the risk assessment requires refinement. Further, a preliminary analysis of a major portion of the avian reproduction studies submitted in support of registration and reregistration shows that no statistically significant effects were found for approximately 50 % of the studies. Thus, I would



suggest that no refinement would be necessary greater than 50% of the time since there will be cases where effects would be found but at test levels significantly below the estimated environmental concentrations. The analysis also revealed that when effects were found, 60% of the time it was found to be egg production. If this pattern holds in the future, the costs for a dose-response avian reproduction tests for this response would be measurably reduced. The redesigned avian test would require a more thorough analysis of existing data and some modeling to evaluate the utility of the new testing design.

27. Page 4-31. Re: Table 4.4-1 Various sources of variability associated with the slope of the dose-response curve. On page 4-70 it states that the level of variability noted here tends to suggest that inter-species differences do not contribute much more than what is already present. Since the data sets are relatively small, especially within test and among species (not > than 45), it would be advisable to analyze the data in the pesticides EcoTox database to re-evaluate and see if trend still holds.

28. Page 4-34 -35. The current lack of data on age-dependent toxicity of pesticides should not preclude their incorporation into probabilistic risk assessments. Again, I recommend investigation by an expert panel or workshop to determine the importance of this factor in influencing the risk assessment and even provide expert judgement that could be used to develop a distribution for use in probabilistic assessments. Concurrent efforts should be made to develop data that would address age-dependent toxicity.

29. Page 4-48. Some rational and justification needs to be added to support the choice of the fixed level of protection (a level which encompasses 95% of the predicted species sensitivity distribution) and what this level means to an individual, and the population.

30. Page 4-49. Some special note here is needed to remind the reader that the analysis is limited to cholinesterase-inhibiting chemicals. The conclusions may not apply to chemicals with other modes of action. Further analysis with chemicals representing a wide range of modes of action is needed.

31. Page 4-52 and -53. The workgroup members state that "little can be done at this time, with current knowledge, to address [the criticisms]" leveled by some regarding the use of distribution-based extrapolation models. They state however, and I agree, "that the adoption of a distribution approach to dealing with species differences in sensitivity is an improvement to the assessment of risks to wildlife and essential to probabilistic assessments." Since the use of distribution-based extrapolation models is key to the whole approach proposed for a probabilistic effects assessment, I recommend a workshop of experts to discuss and address the criticisms and recommend approaches to evaluate, and characterize the assumptions inherent in these models.

32. Page 4-53 and -54. I believe that the optimum number of species to test for acute LD50 in order to establish the 5<sup>th</sup> percentile needs further work. When a chemical needs refinement, and interspecific variation is important, why not five or six additional species LD50 tests? As noted by the workgroup, four is “somewhat arbitrary” and “for the moment.” I recommend that the support statement be changed to support testing of four or more species. Additional changes also will be required in the following pages.

33. Page 4-55 to -64. RE: Extrapolation Factors to predict a predetermined protection level. The method using historical data developed by Luttik & Aldenberg (1995) assumes that species sensitivities are randomly distributed without any trends or patterns associated with taxonomy, save the exceptions noted on page 4-51 for the Icteridae and the Phasianidae for cholinesterase-inhibiting chemicals. First, I would recommend that additional analysis is needed for other classes of chemicals. Second, my brief analysis of the 171 LD50 studies in the EFED data base reveals a trend: for studies where the probit slope was reported, songbirds have lower LD50 values and steeper slopes. I would also recommend that further analysis of extant data is necessary before we can confidently apply extrapolation factors.

34. The important point on lines 14 and 15 of page 4-54 should be put in **bold**, and the entire paragraph should be copied as a footnote to Table 4.6-2 on page 4-78.

35. Since extrapolations based on analysis of historical data play such an important role in the development of the effects distributions, I recommend that significant effort be directed toward an the QA/QC and analysis our current ecotox data base.

36. Further, I recommend that for the short and medium periods of exposure (as shown in Tables 4.6-2. and -3), at least four additional species be tested beginning at Level II of Refinement for acute avian testing versus waiting until Level III. Avian reproduction testing should be part of the short and medium periods of exposure. As I discussed above, for long periods of exposure (as shown in Table 4.6-4), a dose-response avian reproduction test should be required at Level II so that the data could be used to extrapolate to a distribution. If the focal species is different from a precocial upland game bird or a waterfowl, then extrapolation factors should be developed via a panel or workshop of experts in techniques for situations where few data exist or where knowledge is lacking. These recommendations are in keeping with the notion that once refinement is needed, the only way to reduce uncertainty is with additional data. In addition, pesticides that currently need to go through refined risk assessments are always required to provide additional data. In many situations the registrant voluntarily submits additional laboratory and field testing for both exposure and effects. Risk assessors from both the Agency and the registrants recognize that additional data reduces uncertainty, albeit until recently, perhaps only qualitatively.

37. Page 4-70. Under Option 1, the workgroup recommended to use the one

slope as the mean of a distribution of slopes. There should be a caution here. I have found that the slopes for LD50 tests on songbirds tend to be steeper than those for bobwhite quail, mallard ducks, Japanese quail and ring-necked pheasants (~3.0 versus ~5.0). If risk assessors have information that the one slope may not be representative of the focal species, then they should recommend a higher percentage of variation or a alternate approach to arrive at both the mean and the variation.

38. Page 4-74. I agree with the workgroup's proposal to assume that the interspecies variability seen with the LD50 test is applicable to the LC50 test. However, I would recommend that a Task Force be charged to carry out a limited LC50 testing program with three additional species for a few different pesticides. This could provide the data necessary to test this assumption.

39. Page 4-75. I agree that the extrapolation factors from the avian LD50 work should not be applied to the avian reproduction study. However, I disagree with the conclusion that "for the purpose of probabilistic risk assessment, the reproduction endpoint would always consist of one point." I recommend changes to wording in this section to reflect what I recommended above for page 4-20 to -22. The changes also should be reflected in Table 4.6-4.

40. Page 4-77 to -80. The explanation of the Levels of Refinement for avian toxicity testing as opposed to a Tier system focuses on the iterative nature of the progression through the levels versus a more rigid step-wise process. There is a heavy reliance on sensitivity analysis to indicate the need for refinement in exposure or effects or both. I recommend a Workshop with invited experts to provide guidance to and training for risk assessors on the proper use of this technique.

41. Pages 4-78 to 4-80. The "Uncertainty not accounted for" section in each table should be footnoted. The footnote should say something like the following: "Where few data exist or where knowledge is lacking about these factors, the workgroup has identified certain methods that would be helpful in these situations: subjective or Bayesian statistical methods, employing maximum entropy criteria to select distributions from a priori constraints, focusing on extreme value distributions when the tails are of interest, gathering empirically fitted distributions, and using default distributions such as the triangular or exponential (see pages 1-12 to 1-15)."

42. The recommended changes to the tables in this section should also be incorporated into Table 6.2-1.

#### *Process – Levels of Refinement for the Assessment Process*

43. Section 6.0. The effectiveness of these Levels of Refinement prior to implementation, urgently needs Proof of Concept analysis and research ASAP.

44. Pages 6-7 to 6-9. Levels of Refinement versus Tiers. The flexibility inherent in the Levels of Refinement approach versus the Tier approach provides the opportunity to refine where data is available to refine. However, I have concerns with this approach. Until the case studies are completed to provide proof of concept, there will be little or no consistency. The risk management divisions utilize consistency as a very important factor in their decision-making. Second, I wonder if risk assessors could easily misuse this approach? It would be possible to focus refinement (on either exposure or effects) in order to either to reduce the risk below the acceptable threshold or above it, rather than to reduce the uncertainty of the risk. The latter is the goal of probabilistic risk assessment. It appears that even the workgroup may have fallen into such a trap on page 6-15 when it stated, "If including this in the assessment was not enough to *reduce the risk below threshold*, one would hesitate to proceed with radio tracking." (Italics are mine) Also illustrative is the chlorpyrifos example in Appendix C-10. It seems to show that refining the exposure results in a decrease in the risk estimate, while refining the effects results in increasing the risk estimate.

Thus, for an interim period during implementation, perhaps five years, I recommend that a Tiered approach be adopted. This Tiered system would require relatively equal levels of refinement for both sides of the risk function.

45. Page 6-8 to -14. Threshold of Acceptability. The threshold of acceptability will be very difficult to define. Risk managers alone will not be able to define it. There have been ongoing efforts under the auspices of the EPA Risk Assessment Forum to define what is important to protect. Establishing a threshold for acceptability will take considerably longer. While the fact that risk management decisions are made implies that such thresholds exist, there are other factors not accounted for in this report that contribute to decision-making. Political pressure, benefits to the farmer and the public in general, overriding health effects issues or lack of them, perceptions of risk based on comparisons with other pesticides, are only a few of the factors which can rise or lower a threshold of acceptability. In any case, the 'band' representing the threshold in Figure 6.7-1 needs some bounds beside cost in time and money. I recommend that a workshop be organized where experts representing different endpoints can attempt to establish minimum threshold levels for adverse effects representing local, regional, national, and endangered species populations. Such expert scientific input will help to define the lower bound of the 'band' of acceptability.

#### *Methods & Tools and Process – Recommendations*

46. Section 7.0. I recommend that the workgroup provide some perspective here. First, the result of probabilistic risk assessment process is not the final product given to the risk manager. The probabilistic risk assessment must be weighed along with other non-quantitative information such as incident reports and proximity of use and potential adverse effects to sensitive wildlife habitat including endangered species, etc. Thus, probabilistic risk assessment is just one component, albeit a crucial one, in the weight

of evidence analysis. This analysis is typically a part of the final risk characterization.

47. Page 7-4. On line 21 and 22 the workgroup states that many of the of the techniques for combining information on exposure and effects to characterize risk can be used immediately. A specific list of those techniques should be included here.

48. Pages 7-5 to -21. I agree that analysis of several case studies should be a near-term goal of the implementation plan, and that risk managers should participate in the review of the case studies.

49. Pages 7-15 and -16. In section 7.5, "Process for Carrying Out the Recommendations", the workgroup discouraged the Agency and registrants from attempting to develop the needed information as part of evaluating new chemicals, i.e., on a case-by-case basis. They argued that this is a piecemeal approach and would not permit sufficient standardization of individual activities or coordination of overall programs. They encouraged partnering of Industry, EFED, ORD, and other interested groups as needed. They mentioned developing a Cooperative Research and Development Agreement (CRADA) or and informal research steering committee as mechanism to initiate the projects necessary to develop the information, noting that the former would be very time consuming.

50. I would argue that once the approach to collecting information is standardized, then collecting it on a case-by-case basis is a tried way to build important data bases. One good example is the EcoTox database. However, there are efficiencies to be gained from partnering. Toward this end, I would recommend that the Spray Drift Task Force is a useful model for developing this information. Both the Agency and Industry were wearied by scattered case-by-case attempts to establish parameters for estimating aquatic exposure contributions from spray drift. Thus, a partnership was established to pool public and private resources in order to develop a theoretical model and gather sound measured field data on spray drift. This empirical data was then used to load the model and validate the outputs.

51. I would recommend that a Terrestrial Exposure Task Force and a Terrestrial Effects Task Force be developed. Following the example of the Spray Drift Task Force, each would be charged with gathering sound residue data and effects data to load into the models and tools developed by this workgroup.

52. In the meantime, I also would recommend that case-by-case data be generated via the registration and reregistration process whenever a pesticide needs refinement.

#### *Methods & Tools and Process - Appendices*

53. The workgroup provided various modeling tools, equations, databases and

approaches in the Appendices. They included:

- \* overall risk models such as A2, Description of Pesticide Agro-eco Risk Evaluation Toll (PARET), A3, An Individual-Based Model of Pesticide Ingestion and Mortality in Avian Species (Dixon Model), and C10, Risks to Birds from the Use of Chlorpyrifos on Apples: An Example Using ECOFRAM Approaches,
- \* exposure models - C1, PT- Proportion of Diet Obtained in Treated Area; C2, AV - Avoidance; C3, Granular Exposure Model (GEM) for Birds; C4, Computer Models for Estimating Pesticide Concentrations in Environmental Media (AgDRIFT Spray Drift Model, PRZM3, EXAMS, TEEAM, Compartment Models, UTAB, SNAPS/PLANTX, PLANT, Soil-Plant-Air Fugacity Model, TRIM, Plant Uptake Concentration Factors); C5, Volatilization and Pesticide Concentrations in Air; C6, Pesticide Dissipation Kinetics in Environmental Media;
- \* data requirements and environmental fate databases - C7, U.S. EPA/OPP Required Fate and/or Residue Studies; C8, Environmental Databases, C9, Level 1 and 2 Interim Estimates of Pesticide Concentrations in Environmental Media;
- \* recommendations for the problem formulation phase of the risk assessment process B B1, Problem Formulation; B2, Examples of Agro-Ecological Scenarios; B3, Key Species Selection: Recommended Criteria for the Screening-Level, Hypothetical Birds and Mammals.

Each appendix could be the subject for a workshop. There is no indication that each was the subject of a consensus of experts. Nor was there any comparative evaluation of the models, equations, databases, recommendations, etc. The information appeared to be provided as starting points for the implementation phase of ECOFRAM. Some explanation and evaluation of these appendices would help the readers understand why they are included and how the workgroup intended they be used.

54. Figure 10. In Appendix C10-16, shows higher risk when a refined estimate of the species sensitivity distribution is included. Thus, when additional toxicity data is factored into the assessment, the risk is increased. Thus, might additional toxicity data be a disincentive for probabilistic assessments, especially when the standard test species may prove to be relatively less sensitive than focal species? Rather than providing a clear example of the use of the proposed criteria and methods, this example seemed to raise more questions and add more confusion. Consider another example??

#### Response to Questions identified in the "Charge to the Workshop Panel Members"

1. Is the draft report scientifically sound? If not, please explain and provide specific

suggestions on how to improve the report to make it scientifically sound.

Response: The report, including the methods, tools, models, and approaches, appears to be based on the best scientific information available to the workgroup members, considering their resource limitations. However, the workgroup suggests that assessments would not be scientifically defensible if limitations in the data and/or the understanding of the system requires assumptions which introduce large uncertainty in the predicted effects. (Page 1-19, lines 8-11) I believe that there are many factors and parameters that are not accounted for in the report that may introduce large uncertainty in the predicted effects (e.g., "Uncertainty not accounted for" in Tables 4.6-2, -3, -4). Thus, in my specific comments above, I have attempted to identify those parameters and recommend approaches to establish their importance and contribution to the overall uncertainty in the risk assessment.

2. Did the ECOFRAM Workgroup address the Charge to the Terrestrial and Aquatic Workgroups identified in the background document, Evaluating Ecological Risk: Developing FIFRA Probabilistic Tools and Processes? If not, please explain why not and provide specific suggestions on how the Charge could be addressed.

Response: I specifically responded to this question in paragraphs 1 through 3 above.

3. What are the limitations for predicting risk using the approach described in the draft report? Please provide suggestions.

Response: I refer to my specific comments above. To its credit, the workgroup identified many limitations in the proposed approach. Where the workgroup members identified parameters that it felt were beyond the scope of the charge due constraints of time or available data, I recommended that steps be taken to address the importance and potential uncertainty contributed by these parameters. I felt that since we are moving down the path to provide the magnitude and probability of actual risk, we should not 'a priori' ignore any potentially important parameter that can contribute uncertainty in the final probabilistic assessment result. By moving from deterministic to probabilistic ecological risk assessments we strive to establish better levels of trust between risk assessors and risk managers. This demands that we make every attempt possible to include all sources of variation and uncertainty in these assessments.

4. Taking into account your answers to the three questions above, what areas of the report need to be strengthened? If possible please provide specific recommendations for how to strengthen the report.

Response: I refer to my specific comments above. All areas of the report need to be

strengthened. I focused specifically on the effects assessment and have provided specific recommendations that I believe will strengthen the report findings prior to initiating an implementation phase.

5. At what point in the risk assessment process is the certainty level high enough to support the consideration of risk mitigation? What is the minimum level of technical information and scientific understanding that is necessary to evaluate whether risk mitigation would be necessary and/or effective?

Response: I would argue that the certainty level necessary to support consideration of risk mitigation and the minimum level of technical information and scientific understanding that is necessary to evaluate whether risk mitigation would be necessary and/or effective can be based upon previous risk management decisions, i.e., granular carbofuran, diazinon use on golf courses and sod farms, and azinphos-methyl use in sugarcane. In these cases the certainty was not established quantitatively by probabilistic methods, but qualitatively by the weight of evidence consideration of incident data. Thus for previously registered pesticides, deterministic quotients along with incident information are sufficient to support consideration of risk mitigation. Some expression of the magnitude of effects would be necessary to evaluate whether risk mitigation would be necessary and/or effective. This would mean, for example, that the acute risk would have to be expressed as mortality. Finally, for new chemicals without a significant use history, I believe that probabilistic distributions based on testing focal species or greater than four species as well as residues measured in the field will be required to establish the certainty necessary to support the consideration of risk mitigation.